LECTURE 41

Concentration of Measure

- All concentration inequalities we studied so far (Hoeffding, Chernoff, Bernstein, matrix Bernstein) are for <u>linear</u> functions of r.v's (their sums).

Nonlinear: i.e. $|f(x)-f(y)| \le L \|x-y\|_2 \quad \forall x,y \in \mathbb{R}^d$

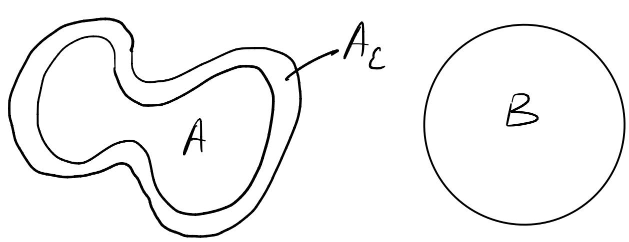
<u>THM</u> (<u>Gaussian concentration</u>)

Let $f: \mathbb{R}^d \to \mathbb{R}$ be an L-Lipschitz function, and $X \sim N(0, I_d)$. Then

$$\mathbb{P}\{ |f(X) - \mathbb{E} f(X)| \ge t \} \le 2 \exp\left(-\frac{t^2}{2L^2} \right) \qquad \forall t \ge 0$$

The proof is based on an <u>isoperimetric inequality</u>:

"Among all subsets of $\mathbb{R}^d$ with given volume, the Euclidean balls have minimal surface area"



More generally, consider the ε-neighborhood of $A \subset \mathbb{R}^d$:

$$A_\varepsilon := \{ x \in \mathbb{R}^d : \exists y \in A \ \|x-y\|_2 \le \varepsilon \}$$

$$\text{Surface area}(A) = \lim_{\varepsilon \downarrow 0} \frac{\mathrm{Vol}(A_\varepsilon) - \mathrm{Vol}(A)}{\varepsilon}$$

<u>Isoperimetric ineq:</u> $\forall \varepsilon > 0$, Among all subsets $A \subset \mathbb{R}^d$ with given volume, the Euclidean balls minimize the volume of A_ε.

Gaussian isoperimetric ineq.: vol $\longmapsto$ Gauss measure

$$\gamma(A) = \mathbb{P}\{\underset{\substack{\uparrow\\ N(0,I_d)}}{X} \in A\} = \frac{1}{(2\pi)^{d/2}} \int_A e^{-\|x\|_2^2/2}\, dx$$

THM (Gauss I.E.)

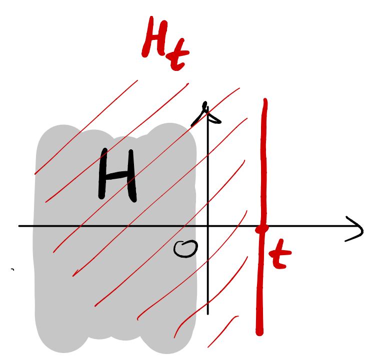
> Let $\varepsilon > 0$. Among all sets $A \subset \mathbb{R}^d$ with given Gauss. measure, the half-spaces minimize the Gauss measure of A_ε.

- Hence: $\gamma(A) \geq \gamma(H)$ implies $\gamma(A_\varepsilon) \geq \gamma(H_\varepsilon)$ $\forall$ half-space H. $(*)$

Proof of Gauss. concentration Thm (p.1): WLOG $L = 1$.

- Let M = median of $f(X)$, i.e.

$$\mathbb{P}\{f(X) \leq M\} \geq 1/2, \qquad \mathbb{P}\{f(X) \geq M\} \geq 1/2.$$



- Let $A := \{x \in \mathbb{R}^d : f(x) \leq M\}$, $H := \{x \in \mathbb{R}^d : x_1 \leq 0\}$.

$$\gamma(A) \geq \frac{1}{2} = \gamma(H) \overset{(*)}{\Longrightarrow}$$

$$\gamma(A_t) \geq \gamma(\underset{\substack{\| \\ \{x\,:\,x_1 \leq t\}}}{H_t}) = \mathbb{P}\{\underset{\substack{\wr \\ N(0,1)}}{X_1} \leq t\} = 1 - \mathbb{P}\{X_1 > t\} \overset{\text{Gauss. tail}}{\geq} 1 - \exp(-t^2/2)$$

- Claim: $A_t \subset \{x : f(x) \leq M + t\}$

  (Indeed, if $x \in A_t$, then $\exists\, y \in A$: $\|x - y\|_2 \leq t \xRightarrow{f \text{ is } 1\text{-Lip}}$ $|f(x) - f(y)| \leq t \overset{\Delta}{\Rightarrow} f(x) \leq f(y) + t \leq M + t$)

- $\Rightarrow \mathbb{P}\{f(X) \leq M + t\} \geq \mathbb{P}\{X \in A_t\} = \gamma(A_t) \geq 1 - \exp(-t^2/2)$

$$\Rightarrow \mathbb{P}\{f(X) - M \geq t\} \leq \exp(-t^2/2)$$

Repeat for $-f \Rightarrow \mathbb{P}\{|f(X) - M| \geq t\} \leq \exp(-t^2/2)$

- Replacing median by mean: **DIY.**

QED

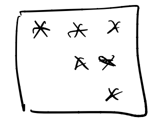
Application to random matrices:

THM (Concentration of eigenvalues) Let A be a $n\times n$ symmetric random matrix with iid $N(0,1)$ entries on & above diagonal. Then, $\forall\, i\in\{1,\dots,n\}$:

$$\mathbb{P}\left\{ |\lambda_i(A) - \mathbb{E}\lambda_i(A)| \geq t \right\} \leq 2\exp\left(-\frac{t^2}{4}\right), \quad t\geq 0$$

Proof

- Vectorize $\Rightarrow$ identify A with the vector $\bar{A}\in\mathbb{R}^{n(n+1)/2}$ that contains the entries of A on & above diagonal.



- Apply Gauss. concentration for $f:\mathbb{R}^{n(n+1)/2}\to\mathbb{R}$:

$$f(\bar{A}) = \lambda_i(A).$$

- $|f(A)-f(B)| = |\lambda_i(A)-\lambda_i(B)| \leq \|A-B\|$ — operator norm (Weyl's inequality)

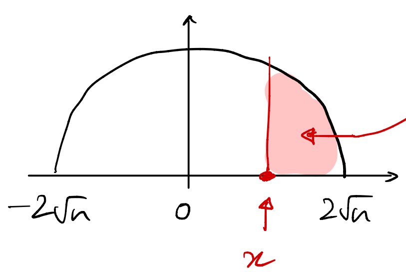
$$\leq \|A-B\|_F \leq \sqrt{2}\,\|\bar{A}-\bar{B}\|_2$$

(square $\Rightarrow$ LHS = sum of n^2 entries, RHS = 2·sum of entries on & above diagonal.)

$\Rightarrow f$ is $\sqrt{2}$-Lipschitz.

- Gauss. concentration Thm $\Rightarrow$ QED.

Q: $\mathbb{E}\lambda_i(A) = ?$ Use Wigner's semicircle law (lec. 25, Oct. 31):



area $\simeq \frac{1}{n}\cdot\#(\text{eigenvalues} \geq x)$

Set area $:= i/n$, solve for $x = \mathbb{E}\lambda_i(A)$.

CONCLUDING REMARKS

Further literature:

1. My book (e.g. beyond Gauss. concentration)
2. Joel Tropp "Probability in high dimensions" ✓
3. Van Handel "Probability in high dimensions" — more advanced ✓
4. Wainwright "High-dimensional statistics" ✓

2ND edition